

8 1999

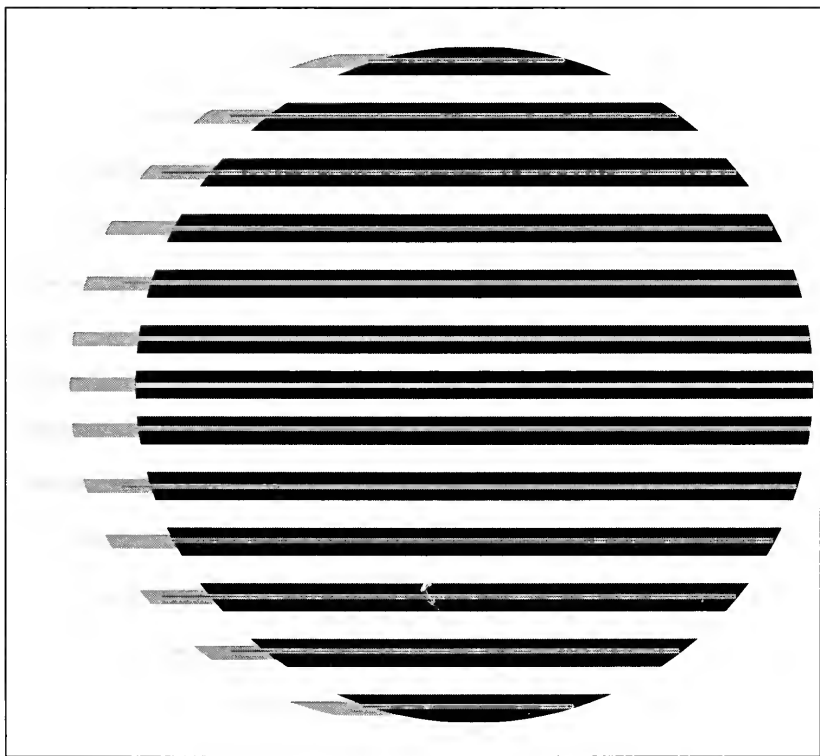
IASSIST

Q U A R T E R L Y

VOLUME 22

Winter 1998

NUMBER 4



Digitized by the Internet Archive
in 2010 with funding from
University of North Carolina at Chapel Hill

IASSIST QUARTERLY

The IASSIST QUARTERLY represents an international cooperative effort on the part of individuals managing, operating, or using machine-readable data archives, data libraries, and data services. The QUARTERLY reports on activities related to the production, acquisition, preservation, processing, distribution, and use of machine-readable data carried out by its members and others in the international social science community. Your contributions and suggestions for topics of interest are welcomed. The views set forth by authors of articles contained in this publication are not necessarily those of IASSIST.

Information for Authors:

The QUARTERLY is published four times per year. Authors are encouraged to submit papers as word processing files. Hard copy submissions may be required in some instances. Word processing files may be sent via email to jstratford@ucdavis.edu. Manuscripts should be sent to Editor: Juri Stratford, Government Information and Maps Department, Shields Library, University of California, 100 North West Quad, Davis, California 95616-5292. Phone: (530) 752-1624.

The first page should contain the article title, author's name, affiliation, address to which correspondence may be sent, and telephone number. Footnotes and bibliographic citations should be consistent in style, preferably following a standard authority such as the University of Chicago press *Manual of Style* or Kate L. Turabian's *Manual for Writers*. Where appropriate, machine-readable data files should be cited with bibliographic citations consistent in style with Dodd, Sue A. "Bibliographic references for numeric social science data files: suggested guidelines". *Journal of the American Society for Information Science* 30(2):77-82, March 1979. Announcements of conferences, training sessions, or the like, are welcomed and should include a mailing address and a telephone number for the director of the event or for the organization sponsoring the event.

Editors

Karsten Boye Rasmussen,
Eckersbergsvej 56,
5230 Odense M,
Denmark.
Phone: +45 6612 9811,
Email: kbr@sam.sdu.dk

Juri Stratford
Government Information and
Maps Department,
Shields Library,
University of California,
100 North West Quad,
Davis, California 95616-5292
Phone: (530) 752-1624.
Email: jstratford@ucdavis.edu

Production

Laura Bartolo,
Libraries and Media
Services,
Kent State University,
Ohio 44242.
Phone: (330) 672-3024, x31.
Email:
lbartolo@kentvm.kent.edu

Walter Piovesan
Research Data Library
Simon Fraser University
Burnaby, B.C.
Canada V5A 1S6.
Phone: (604) 291-5937.
Email: walter@sfu.ca

Title: Newsletter - International Association for Social
Science Information Service and Technology

ISSN - United States: 0739-1137 © 1998 by IASSIST. All
rights reserved.

C O N T E N T S

Volume 22

Number 4

Winter 1998



FEATURES

- 4 **Global Access and Local Support to the
Processes of European Integration in Central
and Eastern Europe Through Global
Networking**
Dusan Soltes
- 10 **Can the Library and the Data Archive Meet in
Active Support of Research in the Social
Sciences? The Case of ILSES**
Repke de Vries
- 12 **The Role of the Web in the Provision of
National Data and Information Services: the
MIDAS Experience**
J. Chruszcz, K. Cole, and A. McCombe
- 17 **The History Data Service - Using Technology
to Enhance Access**
C. Chappell, O. Struijve, and S. Anderson

Global Access and Local Support to the Processes of European Integration in Central and Eastern Europe Through Global Networking

INTRODUCTION

One of the dominant features of the overall socio-economic development since the end of the Second World War and especially exhilarating one during the last decades has been the process of European integration. The process which started as a relatively humble goal of preventing for the future any new devastating wars in Europe gradually has become one of the dominant factors not only of the whole development in Europe but to some extent also worldwide and one of the best examples of the gradual regionalization and globalization of the contemporary world which has become to some extent a model emulated all over the world by various regional and sub-regional clusters of countries.

In the next parts of this paper we will deal in more details with some specific features of this process in the context of the contemporary development trends in the European integration vis-a-vis the processes of the EU enlargement to the countries of Central and Eastern Europe and especially with the role of the modern information technologies in the support of these processes.

PRESENT STATUS OF THE EUROPEAN UNION DEVELOPMENT AND THE CHALLENGES OF ITS ENLARGEMENT TO CENTRAL AND EASTERN EUROPE

During the last over fifty years of its existence, the European Union (EU) has been passing through the development process which could be characterized by at least two main characteristics. They have been as follows:

- a) gradual and steady enlargement of the EU
- b) systematic increase and development of the common institutions, legislation, various rules and regulations of common policies and all of them leading to further accelerated development but also strengthening of the EU as a whole and its member states individually. In case of necessity, the development of individual relatively weaker member states has been supported through various programs and funds of particular common policies and funds (agriculture, cohesion, infrastructure, etc.)

by Dusan Soltes *

a) In this respect, the EU has developed from its original six members in the mid of 1950s (France, Belgium, the Netherlands, Luxembourg, Italy and Germany) through gradual enlargements to the existing 15 members Union. In addition to the original six members, the present EU includes another nine members who joined the Union later on

as the United Kingdom, Ireland, Denmark in 1970s, Greece, Spain, Portugal in 1980s and Austria, Finland, Sweden in 1990s.

In view of this its gradual development, the EU has become one of the strongest economic powers in the world and in many respects even the most strongest at all. This strength has been increasing not only by each new member but also by the synergic effect of their mutual cooperation and integration

b) all above processes of mutual cooperation and integration in no way should be understood as a simple consequence of the gradual process of enlargement of the EU covering at present - with few exemptions (Norway, Switzerland) the most advanced part of the continent. At least as important as the enlargement itself and to some extent even more important has been the gradual development of the institutions, legislation and various other common policies, funds, rules and regulations of the EU which in their mutual interaction have further substantially contributed to the acceleration of the overall development of the Union. In this respect we could come to the conclusion that in many ways the EU has gradually developed the unique system of international institutions, legislation, "law", etc. which of course has to be fully respected not only by all member states but also by all partners outside the EU and in particular by all candidates for the future membership.

In this respect the most important institutions of the EU which have been overseeing the overall development of the Union and in this respect also being institutions producing the particular rules and regulations of the common policies have been:

- the European Commission

- the European Council
- the European Parliament
- the European Court of Justice

These basic four institutions of the EU are further supported by a large number of various other, more specialized and/or technically oriented institutions, agencies, etc. Altogether, it represents a staff of more than 20,000 international civil servants serving in three main hubs of the EU i.e. Brussels, Luxembourg, Strassburg and having their official representatives and missions all over the world. Of these over 20,000 staff, more than 16,000 have been working directly at the European Commission at Brussels as the most important and powerful executive arm of the EU providing and executing the day-to-day functioning of the EU internally but also externally and sometimes unjustly referred as "bureaucrats" or "euro-bureaucracy" due to the enormous number of various legislation, rules and regulations they have been permanently producing for every aspect of the community life and activities.

If in view of the above we would try at least very briefly to assess the challenges of the current process of enlargement of the EU to the Central and Eastern Europe we have to take into account several important aspects. All candidate countries for this forthcoming enlargement are former socialist countries which emerged in the end of the 1980s as new future democracies and market economies. Since that time on both sides i.e. the EU as well as these new democracies it was clear that the future development of the EU has to proceed towards its further enlargement by these ten new candidate countries if the Europe would like to eliminate any negative consequences of its further and/or continuing division. However, due to the completely different political and socio-economic development in both parts of the Europe i.e. in the EU itself on the one side and in the candidate countries on the other, the process of unification and integration of the whole Europe has been much more complex, time consuming, fund demanding, etc. than it had been assumed in the end of 1980s and beginning of 1990s when the necessary political preconditions for this kind of integration have been created.

Among various other differences as e.g. much lower level of the overall development in the candidate countries (only about 6% of the GDP of the EU while by the population it is more than 1/3 of its current population), the main problem is the necessity to prepare the candidate countries for their future in the EU in the terms of legislation and all various community rules and regulation. What in general has been named as a process of approximation and harmonization of legislation of the candidate countries to the so-called "acquis communautaire" of the EU or its community legislation in its broadest sense.

ROLE OF GLOBAL NETWORKING IN THE PROCESS OF THE FUTURE ENLARGEMENT OF THE EUROPEAN UNION

In order to understand the complexity of the above processes of the approximation and harmonization of the legislation of the candidate countries with the totally different legislation from their former "socialist" past with the "acquis" we have to realize that the later one represents an enormous volume of various legislative acts which in general consists of two main parts:

- primary legislation - all basic treaties by which the EU and before that the European Communities have been established i.e. the Paris Treaty, two Rome treaties, Maastricht and Amsterdam treaties respectively represent the basic part of the "acquis" to which every new member has to access in full and without any exemptions. That of course requires from them to have also the suitable technical tools and means for such an access
- secondary legislation - an enormous amount of various legislative acts, norms, regulations and standards permanently enacted by the EU for the day-to-day running of the EU and mostly initiated by the European Commission. In this case again the candidate countries have to adopt and implement all of them in their national legislation. But in difference to the primary legislation, in this case with the possibility of some flexibility in adoption regarding some of them. What on one hand makes the process of harmonization and approximation more flexible but on the other hand it is also more challenging and more complex in seeking the most optimal ways of the particular adaptation according to the national needs and priorities. But in any case it again requires an efficient and direct access to the particular sources of the secondary legislation.

In view of the above it is clear that the whole process is very closely related to the efficient utilization of the modern contemporary information and communication technologies and their global networking. It has been a historical coincidence that the processes of European integration in the Central and Eastern Europe have started and been proceeding in the environment of the acceleration of the global networking and direct access to the remote data bases and at the same time of a possibility to use networking also for the local support to these processes. As we will demonstrate in the following parts of this paper, some activities in this respect would almost be impossible if this particular "global" access and "local" support through the contemporary networking technologies would be not existing.

In this connection we have to realize that during its over fifty year existence and mainly due to its relatively extensive institutional framework (in particular regarding

the European Commission), the EU has up to now accumulated an enormous amount of "acquis" and related and/or derived or supporting information, documents, acts, etc. According to the latest available edition it represents a list of:

- 51 specialized data bases of what 46 have been on-line data bases and 5 CD-ROMs

- 4 World Wide Web servers

- in addition there have been a long list and ever growing number of other specialized on-line and off-line data bases, new editions of CD-ROMs and www pages e.g. in the TAIEX Office of the European Commission as an office created in 1995 for providing technical assistance to the candidate countries in approximation of legislation according to the White Book but also in mastering the latest information and communications technologies.

In addition we have to realize that the amount of these data sources and data bases has been further increased by the fact that in accordance with the community rules all of them are either in all 11 official languages of the EU (every member country's language is an official language of the EU) or in some of them (mostly English, French, German) with at least annotations in all other.

The list of the EU data bases is as follows:

ABEL - Document delivery of Official Journal L and C series

AGREP - Agricultural research projects of the EU

APC - Commission preparatory acts

BACH - Harmonized company acts

CCL-TRAIN - Common command language training database

CELEX - Community legislation, case-law, preparatory acts, parliamentary questions, national provisions implementing directives

COMEXT - Intra and extra EU trade (also on CD ROM)

CORDIS - Research and development information service (also CD ROM)

CORDIS RTD ACRONYMS - Research and technology development

CORDIS RTD-COM DOCUMENTS - Commission's initiatives

CORDIS RTD-CONTACTS - Contact points of Cordis database

CORDIS RTD-EOI - Expressions of interest

CORDIS RTD-NEWS - Latest news on RTD

CORDIS RTD-PARTNERS - Partner search service

CORDIS RTD-PROGRAMMES - EU-funded research programmes

CORDIS RTD-PROJECTS - Details of projects

CORDIS RTD-PUBLICATIONS - Abstracts of EU publications

CORDIS RTD-RESULTS - Information on results and prototypes

ECDIN - Environmental chemical data (EU, USA, Japan)

ECHO NEWS - Echo news for users

ECLAS - European Commission library

ECU - European currency unit

EMIRE - European employment and industrial relations

EPISTEL - European Parliament press information system

EPOQUE - European Parliament on-line query system

EUROHISTAR - European historical archives

EURISTOTE - Academic research on European integration

EUROCRON - General European Union statistics

EURODICAUTOM - Directory of terminology

EUROFARM CD-ROM - Statistics on agricultural holdings

EUROLIB-PER - Collective catalogue of periodicals

EUROSTAT CD-ROM - Electronic statistical yearbook of the EU

HTCOR-DB - High-Temperature database

HTM-DB - High temperature materials database

I&T Magazine - Industry, telecoms and information market

I'M GUIDE - Information market guide

INFO 92 - European internal market and its social dimension

IUCLID - Classification and evaluation of existing substances

NEW CRONOS - Macroeconomic statistical database

OIL - Weekly oil bulletin

OVIDE - Information service of the European Parliament

PANORAMA CD-ROM - Panorama of EU industry

RAPID - Up-to-date information on EU activities

REGIO - Regional statistics

REM - Radioactivity environmental monitoring

SCAD - Community documentation access system

SESAME - Energy technology research projects

TED - Tenders electronic daily

THESAURI - Structured vocabularies

TIDE - Technology initiative for disabled and elderly people

Although not all of the above databases have the same importance for the functioning of the EU, its member states and/or candidate countries it is evident that their proper utilization can in many aspects contribute to the better knowledge and/or communication with the EU and its institutions. Especially important it is in the case of the candidate countries as on the bases of some of the above databases they can directly be preparing for the challenges of the future membership. In this respect one of the most important databases is the CELEX which contains "interinstitutional" documentation system for Community law. Not only by this its orientation but even more by its content it is the most valuable source of information on "acquis" as in its subsystems it contains information on:

- Legislation (primary (treaties), secondary,

supplementary)

- Commission proposals

- European Parliament resolutions

- Economic and Social Committee opinions

- Court of Auditors opinions

- Judgments and orders

- Opinions of Advocate-General

- Written parliamentary questions

- Oral questions

- Questions at Questions Time

- Parliament documents

This database is on-line but available against payment only and at present it contains about 200,000 entries with a growth rate of about 10,000 entries per year. The biggest advantage of the CELEX is the fact that it contains not only all kinds of references and accompanying information but also the official full text of all basic legal documents including primary and secondary legislation.

It is evident that such an amount of information could not be effectively handled without possibilities for a global access through a global network. This is the only way how this, the most important source on the community legislation i.e. the back-bone of the whole EU can be made to be directly accessible from the member countries of the EU as well as candidate countries of the Central and Eastern Europe. In this connection we have to realize that this global access is needed not only from the point of view of government institutions but also by any business entity and/or a person being in need to proceed and/or familiarize with the community legislation.

An another important aspect of the above databases from the users point of view is the fact that many of them (directly operated by the EC) are available free of charge while some other are available only on the commercial basis.

GLOBAL ACCESS AND LOCAL SUPPORT TO THE PROCESSES OF EUROPEAN INTEGRATION IN SLOVAKIA AS A CANDIDATE COUNTRY

The above system of the EU databases is physically located mostly at the European Commission at Brussels or in some other institutions of the EU or have been operated by an official database operator under the special arrangements with the European Commission. In order to be accessible

also from the other end e.g. the candidate countries it is necessary to make the necessary arrangements and preparations also in the particular country.

In the specific conditions of the Slovak Republic these preparations have started already in 1995 when the European Agreement on Association with the EU entered into force and it has become clear that it will be necessary to create all necessary preconditions for developing an efficient computer network which would enable all government institutions an access not only to the above EU databases but also it would create conditions for the local support and cooperation between individual government institutions in implementation of the Europe Agreement and one of its most important task i.e. approximation and harmonization of legislation.

On the basis of the technical assistance from the EU, the particular computer network has been established under the PHARE program in the form of the computer network linking together all government departments into a computer network.

This specialized computer network for the processes of European integration has the following main features:

- it has been operated as a specialized network within the broader "GOVNET" i.e. the Governmental network system

- it consists of two hubs served by two servers. One being indicated as "CS" i.e. a central server for the network of 28 computers distributed to all central organs and departments of the state administration. The other one "IAP" i.e. a central server for a network of 12 computers at the Institute of Approximation of Laws (IAL) as a specialized institution of the Government for methodical and technical coordination of the specific processes of the approximation and harmonization of the legislation of the Slovak Republic with the legislation of the EU

- in order to reduce operational costs of an on-line access to the particular data bases in Brussels, they have been physically directly available also in the IAL in the form of CD ROMs with monthly updates for local access and utilization

- the whole system of GOVNET enables through the

Internet facilities a direct connection to the outside world especially to the institutions of the European Union and in particular to the European Commission in Brussels as well as to all other institutions of the EU and/or also other partner international as e.g. OECD in Paris, etc.

The main contributions in addition to the above global access to the information sources of the EU are mainly in the area of the local support to the processes of European integration through the particular computer network. As all the government departments have been linked together to the particular "Intranet" its main benefits are as follows:

- direct access and support to all various specialized databases as they have been created or will be created in support of processes of European integration

- direct support to the processes of approximation and harmonization of laws. Especially it is beneficial in case of the legislative acts having "cross-departmental" character i.e. the responsibility for their approximation lies with several departments. In such cases one of them has been acting as a coordinating "gazetteer" department for all other departments participating in the adaptation process. In particular in these cases the function of the global network as well as a local support is inevitable as only the network can create a working environment for on-line cooperation and joint activities on the same "piece" of legislation at the same time by several departments

- another important contribution of the global access and local support of the network is in the enormous amount of activities related the processes of translation of the "acquis" from the official language of the EU (the Slovak Republic in this case uses English version of the legislation) to the Slovak language and then after its harmonization and approximation back being translated to English for the needs of notification and monitoring by the EU. The whole translation itself without the global access and local support of the network would be almost impossible if we take into account the size of the "acquis" and still relatively low level of English comprehension in the countries of CEEC and Slovakia in particular as a new country which before its independence had only very limited opportunities for international relations, etc.

- in view of this, an important role has been played by the computerized thesauri, key words, glossaries, etc. as prepared by the European Commission for the needs of the candidate countries

- another important aspect of the local support to the approximation and harmonization of laws has been in the area of the global access and local support to the

processes of securing proper interpretation of the approximated, harmonized legislation which has to have exactly the same meaning as the official version of the same legislation of the EU.

- the importance of providing this "common" interpretation of the national legislation with the "acquis" has also been directly supported by the network as the whole process in full utilizes its functions and services. The particular "certificate on the compatibility" when issued by the IAL as the official certification authority of the Government of the Slovak Republic is again available at the network for future use in order to prevent any uncertainties in this respect

- the whole process of the harmonization, approximation of the legislation after being officially certified by the IAL has to be notified to the European Commission and in particular to its specialized TAIEX (Technical Assistance Information Exchange) Office in Brussels which provides not only technical assistance to the candidate countries in the harmonization of legislation but is also administering the particular data base on the progress achieved in the whole process. The particular data base monitors the development and progress achieved in the harmonization of legislation in all ten candidate countries (Estonia, Lithuania, Latvia, Poland, Slovakia, Czech, Hungary, Slovenia, Bulgaria, Romania) on the basis of so-called "Harmonogramme" which traces the whole process of harmonization and serves directly to the evaluation of the particular candidate countries and their process of "approximation" to the EU in this single most important area i.e. legislation.

- of course that the global networking plays a very important role also in many other related areas of the European integration, in particular regarding e.g. the processes of preparation and in-service training of the civil servants for their new tasks vis-a-vis their new responsibilities in the processes of European integration, approximation of laws, implementation of Europe agreement, etc. The particular network in this respect serves as an indispensable source of particular information but at the same time also as a set of tools and means for in service distance education and learning

- the same function has been played also in relation to the university education of the students specializing in European integration as e.g. at the Faculty of Management of the Comenius University in Bratislava as well as at other faculties of the same university e.g. at the Faculty of Law but also at some other faculties not only in Slovakia but also in other candidate countries of CEEC.

CONCLUSION

Also from these few examples it is quite evident that the role of the global networking and its functions in global access and local support to the processes of European integration in the countries of the Central and Eastern Europe are inevitable. They make the whole process of integration much more efficient and in many cases it is not at all an exaggeration if we say that without these functions some activities would be almost impossible to carry out within any reasonable time period. Otherwise, the process of the European integration and unification would be even more time consuming and hard-to-be completed if we realize that in spite of these enormous support from the contemporary information technologies the gap between the countries of the EU and the candidate countries is still rather widening than narrowing. But this issue is already beyond the scope of this paper.

REFERENCES:

Europe Agreement on Association between the European Communities and the Slovak Republic, Slovak Chamber of Trade and Industries, Bratislava 1994

Preparation of Associated Countries of Central and Eastern Europe for Integration to the Internal Market of the Union, Volume 1 and 2, The Commission of the European Communities, Brussels 1995

Agenda 2000 - Volume I - Communications: For a Stronger and Wider Union, The Commission of the European Communities, Strassburg 1997

European Union Database Directory, Office for Official Publications of the European Communities, ECSC-EC-EAEC Brussels, Luxembourg 1995

Soltes, D.: Enlargement of the European Union: Some General Trend but also Current Specifics and Peculiarities, CEPS Liechtenstein Conference'97, Brussels, Vaduz 1997

*Paper presented at the 1998 IASSIST/CSS Conference, Yale University, New Haven, USA, May 19-22. 1998.
Dusan Soltes, Faculty of Management, Comenius University, Odbojarov str. 10, Bratislava, Slovakia. Ph:421-7-5666 702 Fax:421-7-5666 703 E-mail:dusan.soltes@mail.fm.uniba.sk

Can the Library and the Data Archive Meet in Active Support of Research in the Social Sciences? *The Case of ILSES*

ILSES: development of tools for an Integrated Library and (Survey) Data Extraction Service. <http://www.gamma.rug.nl/ilses>

Project under the European (EC) Telematics for Libraries Program.

Partners: ProGamma (Netherlands), ZA (Germany), NlWI (Netherlands), University of Amsterdam (Netherlands) and associate partners BSDP (France) and Trinity College (Ireland).

Introduction

Data material collected for empirical research has traditionally been computer stored and electronically distributed by data archives and data libraries. Whereas publications from the same research were kept, referenced and given access to by libraries.

As content providers data archives could not extend their services with relevant book and journal collections, cross referencing and lending of printed material. Libraries could not give access to data related to published research or had the means to expand bibliographic references to also point at data as machine readable outcome of the research process.

A situation where data and books are separately referenced without consistent cross linking, have to be searched for in separate catalogues and are given access to by different authorities and with different facilities, has consequences for any one embarking upon new research or in general needing social scientific information. It is not possible to start with general literature searches in libraries and easily trace back publications to the empirical research and collected data that is at the heart of it. Neither can data archive catalogues (even when expanded with bibliographies) help with book and article searches starting from particular data collecting efforts. Properly linking data and publications would need metadata standards that take such relationships into account and coordinated efforts between authors (proper citation of data sources or writing such metadata directly themselves), the library world (referencing with cross linking in new metadata formats) and the data archives (likewise referencing with cross linking). Part of those efforts would also have to be a

*by Repke de Vries **

common catalogue search facility or some form of easy access from one catalogue to information in the other.

World Wide Web techniques for linking electronic resources on the Internet but also new metadata initiatives that explicitly hold linking information to related (electronic) resources, have the

potential to finally bring data and book together again for searching and retrieval.

A recent publication¹ is referred to for a more complete treatment, including a few Internet related projects that already demonstrate first attempts in this direction. One of these is ICPSR's "Publication Related Archive"², another the European NESSTAR project³. In the same publication ILSES as Integrated Library and Survey-Data Extraction Service, a system of tools and (Internet) facilities, is expanded upon as a current project funded within the Library Programme of the European Commission. ILSES addresses the same goal of integrating publication and data. To achieve this, it accommodates both content providers (libraries and data archives) and end-users.⁴

Other approaches and further developments

In the UK e(lectronic)Lib(raries) program, the Open Journal project has been working on mechanisms and demonstrators for "citation linking" in the broadest sense.

"Using citations - the links made by authors themselves - users can navigate between their current work and a priori work in the archives of the research literature or take a recent paper and move forward, tracking the citations dynamically"⁵

In particular did the Open Journal develop Internet solutions to create, maintain and give access to "distributed links" between primary and subsequent secondary sources, when electronic information is available that never received embedded links or that simply does not have the internal format to adopt such links.⁶ These could easily include formats like data material distributed over the Web, collections following from digitization or MARC type of bibliographic references, which by themselves do not have entries for cross linking. Another advantage of the Open Journal approach could be the fact that information to be

linked in "citation style" fashion, can be arbitrarily distributed over the Net and that establishing the links can be a separate, dedicated activity at any point in time.

Central thesaurus facilities over the Internet are another key issue in tying together distributed information. When metadata can be attached to both data and related publications that take indexing terms from a central (domain specific) thesaurus, future searches will bring up related material because of such common terms. The UK Data Archive has established such an Internet based thesaurus facility, that also takes into account web based forms to submit new possible additions to the thesaurus.⁷

Metadata

The Dublin Core metadata definition is both finalizing in details but also has a core set that already finds application in sometimes large scale projects. Last years DC 5 conference in Helsinki reflect this. Both a series of current DC projects were discussed⁸ and several DC elements received further clarification and more precise definition.⁹ Two projects in particular address the same issue of integrating diverse but related information types - one by the Australian Geodynamics Cooperative Research Centre and one by the UC Berkeley Digital Library Catalog.¹⁰ It is interesting to see the integration approached in a distributed DC metadata fashion instead of by a central database model.

At and following the conference the DC.Relation element (among others) received further definition. The now proposed six sub-elements have enormous potential for cross-referencing and thus having build-in links to go from a DC data material description to following DC descriptions of article and book publications - especially where these are electronically available as well.¹¹

Conclusion

The technology, connectivity and development of standards is available enough to start bridging the two infrastructures of access to data and access to publications following analysis of those data or touching upon the same theme.

ILSES addresses that goal with very concrete tools and solutions. Hopefully it can both be useful for end-users and at the same serve as an evaluation of the type of approach chosen, i.e. a central database keeping all the metadata and the linking information.

It seems that the library world with its digital library projects and its developing metadata standards and distributed models have an advantage for future solutions and more momentum to realize these. The data archiving world should be quick to take up the challenge and start working with library people on the common cause of giving researchers complete access to all the related information of any type, following from their research

activities. From the beginning should this access be complemented by facilities for these same researchers (data collectors, authors, publishers, depositors) to create metadata and linking (citation-) information equally well themselves.

References

1. Vries, R.E.de: ILSES: how library and data archive meet in active support of research in the social sciences. INSPEL Vol. 31 No. 4, 1997. The article is also electronically available in PDF format: <<http://www.fh-potsdam.de/~IFLA/INSPEL>>
 2. ICPSR "Publication Related Archive" <http://www.icpsr.umich.edu/ICPSR/Other_Resources/prs.html>
 3. NESSTAR (Networking European Social Science Tools and Resources) <<http://dawwww.essex.ac.uk/projects/nesstar/>> Though NESSTAR's first goal is integrated access across holdings at different data archives, the choice for Z39.50 at least opens their catalog searching to libraries and thus bridges the separate infrastructures
 4. ILSES <<http://www.gamma.rug.nl/ilses>>
 5. From the IRISS'98 conference: Hitchcock, Steve et al.: Webs of research: putting the user in control. <<http://sosig.ac.uk/iriss/papers/paper42.htm>>
 6. Carr, Leslie et al.: The distributed link service: a tool for publishers, authors and readers. <<http://www.w3.org/Conferences/WWW4/papers/178/>> Please note the IRISS'98 paper in note 5 has references to further work on these tools by the Open Journal Project
 7. The HASSET thesaurus project at the UK Data Archive, jointly with other parties: <<http://biron.essex.ac.uk/cgi-bin/zhasset>>
 8. DC conference October 1997: Project Presentations <<http://linnea.helsinki.fi/meta/projects.html>>
 9. A very thorough account is given by Diann Rusch-Feja: Entwicklungen der Dublin Core Metadaten: Bericht ueber den 5. Dublin Core Metadata Workshop (full title abbreviated) <<http://www.mpib-berlin.mpg.de/DOK/dc5ber.htm>>
 10. Project Presentations 19 and 25 from the document in note 8
 11. See Rusch-Feja's article in note 9 ; the relevant section is almost self explanatory and a series of further URL's of mostly English information sources is given.
- * Paper presented at the 1998 IASSIST/CSS Conference, Yale University, New Haven, USA, May 19-22, 1998, Repke de Vries, ILSES - NIWI: repke.de.vries@niwi.knaw.nl

The Role of the Web in the Provision of National Data and Information Services: *the MIDAS Experience*

MIDAS is a JISC designated national data centre for the UK higher education community providing on-line access and support for a range of large and complex datasets, such as censuses, surveys, time series databanks, bibliographic and full text databases. In this context, MIDAS is part of the developing JISC funded National Distributed Electronic Resource which is seeking to promote and extend access to electronic information and services to the entire UK higher education community.

The expectations of users have changed considerably and we have had to rethink how we deliver data and information to the researcher's desktop. It is not enough to promote awareness of the data resources and their potential applications in teaching and research. We also have to convince the users that their time is being used efficiently, that they can easily identify the data that they want, extract it and where appropriate put it into a suitable format for secondary analysis. For us this means creating appropriate interfaces for the data - simple enough for a once-off selection and versatile enough for more sophisticated use.

This paper addresses the influence of the Web and the expectations of its users on the services provided by MIDAS. It shall describe some of the new interfaces to data and information which will be of particular interest in both research and teaching.

MIDAS overview

MIDAS (<http://www.mimas.ac.uk/>) is a JISC (<http://www.jisc.ac.uk>) designated national data centre for the UK higher education community providing on-line access and support for a range of data and information resources. Along with the other JISC funded data centres (BIDS and EDINA), data services and projects, MIDAS is part of the developing Distributed National Electronic Resource (DNER) which is seeking to promote and extend access to electronic information and services to the entire UK higher education community (http://www.jisc.ac.uk/ccr/dner_colpol.html). MIDAS is based at Manchester Computing at the University of Manchester. Manchester Computing also hosts other projects, such as COPAC (<http://copac.ac.uk/copac/>) and the SuperJournal Project (<http://www.superjournal.ac.uk/sj/>).

*by Julia Chruszcz, Keith Cole and
Anne McCombe **

The data and information resources available through MIDAS include on-line access to the UK Census of Population; government and other continuous surveys; national and international time series databanks; digital map data; satellite images; chemical information systems; bibliographic data resources and electronic journals. A key component of

the service is the provision of a range of specialist support services, such as documentation, training and research support, relating to the data and information resources available on MIDAS. Many of the services are run in collaboration with other organisations, such as The Data Archive at the University of Essex.

MIDAS also provides access to a range of software packages and the large scale computing resources (i.e. memory, disk space and CPU time) required by users wishing to undertake complex data analysis.

MIDAS also provides facilities and support to projects wishing to exploit the internet to provide wider network access to data and/or information resources for teaching and research purposes. Two key data sharing and gateway services that have developed considerably over the past couple of years include NetEc (<http://netec.mcc.ac.uk>), which is a collection of projects which aim to improve the usefulness of electronic networks in Economics, and GENUKI (<http://www.genuki.org.uk>), which provides information relating to the study of genealogy and family history.

The service continues to grow both in terms of the range of services offered and numbers of registered and active users.

The history of MIDAS and the Web

Historically, Manchester Computing has always made efforts to ensure that service specific information is available on-line. Prior to the installation of a UNIX platform for the MIDAS service in 1993, the on-line help systems were developed using proprietary tools and access was restricted to users logged onto the system. With the advent of a UNIX based service, the on-line information system moved to gopher. This client-server approach to providing access to information about the various datasets, software and other services available via MIDAS proved

extremely flexible. It could be used by users actually logged onto MIDAS as well as users with local access to a gopher client.

Until relatively recently, the gopher server was used as the primary method of providing access to information. With the advent of the Web, the initial function of the MIDAS home page was to simply provide an alternative interface to the information held in the gopher. However, the MIDAS Web site (<http://www.mimas.ac.uk/>) soon developed into a much more sophisticated information resource which was more suited to meeting the diverse information needs of data set users. Service specific Web pages were established; documents, reports and newsletters were made available in an HTML format; experimental Web interfaces to datasets, and information were also developed together with the provision of links to related Web based resources. Compared to the graphical user interface offered by the Web, the hierarchical text based gopher system soon started to look antiquated. As use of the Web and access to Web browsers became more widespread, it seemed appropriate to transfer the MIDAS information server from gopher to Web. Increasingly, the service specific MIDAS Web pages are starting to become information gateways in their own right. A good example are the Web pages relating to the statistical packages on MIDAS (<http://www.mimas.ac.uk/stats/>).

The decision to transfer the information server from gopher to Web was not greeted with universal acclaim. Initial user feedback indicated that whilst the majority of MIDAS users had easy access to Web browsers there was still a sizeable minority of users with only telnet access and/or without access to a Web browser who would be inconvenienced if the gopher service were withdrawn. Consequently, it was necessary to install lynx – the non-graphical Web browser – on the MIDAS UNIX server to provide an alternative method of accessing information held as Web pages on MIDAS.

Responding to changing user expectations

There are currently a variety of different interfaces to the various data and information resources available on MIDAS. In part, this reflects the specialist software requirements of many of the datasets. For example, the majority of the government and other continuous surveys are supplied to MIDAS in SIR database format. Indeed, some surveys, such as the British Household Panel Study (BHPS), are actually supplied in multiple formats (e.g. SIR, SPSS, SAS and STATA). Although some of the services are entirely Web based (e.g. JSTOR, COPAC and ONS Databank) or use specific client-server software packages (eg Beilstein CrossFire), the majority still require the user to physically log onto a remote UNIX server – either via TELNET or X-Windows – and run one or more application packages.

There are a number of problems with this mode of working. The user is required to have a basic competence with the UNIX operating system, a reasonable prior knowledge of the structure of the data as well as expertise in one or more specialist data extraction and analysis packages. Whilst this might not pose a problem for computer literate and network aware researchers, it does represent a significant barrier to less experienced users, such as undergraduate students, across a broad range of disciplines who might have extremely limited data requirements.

The expectations of users have changed dramatically over the last couple of years and we have had to rethink how we deliver data and information to the researcher's desktop. It is not enough to promote awareness of the data resources and their potential applications in teaching and research. We also have to convince the users that their time is being used efficiently, that they can easily identify the data that they want, extract it and put it into a suitable format for secondary analysis and that they can obtain informed advice on types of use. For us this means creating appropriate interfaces for the data – simple enough for a once-off selection and versatile enough for more sophisticated use.

Over the last couple of years MIDAS has endeavoured to respond to changing user expectations by trying to develop more Web based interfaces to many of the data and information resources. A number of the Web based interfaces to MIDAS services have been developed in house using CGI programming techniques. For example, a Web based interface to the ONS time series databank for the UK has been developed (<http://www.mimas.ac.uk/ons/>) in order to facilitate greater use of the data in teaching. This simple interface permits searching by keyword, browsing of series and saving of extracted series in different formats. It is important to note that many of these Web interfaces represent additional interfaces rather than replacements for existing interfaces requiring a UNIX login. The aim is to provide more appropriate interfaces to new categories of users rather force existing users into new modes of working.

Whilst it is relatively easy to develop simple search and retrieval interfaces to data and information resources – providing it does not require too much restructuring of the underlying datasets – it is much harder to develop more sophisticated interfaces. MIDAS is currently exploring the use of java tools to develop a Web enabled version of the cartographic data visualiser (cdv) which uses scientific visualisation techniques for exploratory spatial data analysis (<http://www.mimas.ac.uk/JANUS/cdv/>). However, building complex Web based interactive data extraction and visualisation systems is currently resource intensive and might not be worthwhile given the advent of commercially supported products. In addition, it is clear from the feedback from users that the ease of use of an

interface is often more important than its functionality.

One of the major brakes on developing more Web based interfaces is the current absence of Web enabled versions of many of the packages currently used on MIDAS for providing access to data. This is changing gradually but perhaps too slowly for most users. For example, the SAS/IntrNet software product (<http://www.sas.com/software/components/intrnet.html>) can be used to develop Web-enabled SAS applications running on a central server. Similarly, ESRI Map Objects Internet Map Server (<http://www.esri.com>) provides a set of tools that can be used for serving maps over the Web. Both of these products will be evaluated for use on MIDAS.

Is accessibility the only constraint?

The Web has a major role to play in improving accessibility to electronic data and information resources. Whilst accessibility is a major problem that requires addressing, it is not the only constraint on more widespread and effective use of data and information resources. For certain datasets, increased availability and improved interfaces have not necessarily resulted in the expected growth in numbers of users that might have been expected. One factor that plays an important role is the general lack of awareness amongst the wider academic community about the availability, content, scope, structure of key data resources combined with a lack of understanding about their potential application in both teaching and research. Although some users may have a good awareness about individual datasets this may not be matched by an equivalent understanding of complementary data resources. In addition, the level and type of use of certain data sets can also be adversely affected by a lack of appropriate secondary analysis skills. For example, an absence of spatial data handling skills frequently acts as a major constraint on the use of key spatial data resources, such as 1991 Census digitised boundaries.

The combined problems of accessibility, awareness and usability are being addressed by the KINDS (<http://www.mimas.ac.uk/kinds/>) project with respect to the spatial data resources available on MIDAS (<http://www.mimas.ac.uk/maps/>). An integrated set of Web based tools have been developed which enable relatively inexperienced users to search, browse and work directly with large and complex spatial data sets, such as the Bartholomew's digital map data for Great Britain. These tools include a spatial search engine; data set browsing; a help system which provides access to a knowledge base of geographical terms and concepts together with mapping/data download facilities.

Feedback from users indicates that the complex registration process for many of the copyright/commercial data sets held on MIDAS does act as a major deterrent to use – particularly for teaching purposes. Although registration is

frequently seen as part of the cost of obtaining free or low cost access to commercially valuable datasets there is a need to negotiate unrestricted academic access to more sample datasets to facilitate the development of teaching and learning materials. If the web is to be used to build virtual learning environments it is important that it is populated with meaningful sample datasets.

Developing sustainable data services

Use of data and information resources in academic teaching and research will inevitably generate a range of derived materials. These derived materials could include derived data sets; program code/algorithms; methodological notes; data quality reports or teaching materials. It is regrettable that large amounts of this value added material are frequently discarded and not recycled back to other users. Whilst funding bodies such as JISC are concerned about the long term preservation of electronic materials that constitute the DNER the issue of building dynamic and sustainable databases of derived material has not received much attention to date.

MIDAS has worked with a number of projects and users wishing to make derived datasets and/or software tools more widely available – particularly where this adds value to the services hosted by MIDAS. This has happened particularly with respect to the 1991 Census area and interaction statistics held on MIDAS. An example is the 1981 and 1991 Census population surfaces and associated access software (<http://census.ac.uk/cdu/surpop/>). The KINDS project is also currently exploring developing databases of derived material relating to the spatial data resources available on MIDAS. On the basis of experience to date, it is clear that the Web offers great potential for building collaborative virtual user communities around key data and information resources. However, there are a considerable number of technical, legal, organisational, quality assurance and cultural issues that require resolution before such dynamic web enabled databases of derived materials can be developed.

Exploiting emerging technologies

Keeping in contact with both existing and potential users of MIDAS services and alerting them to new developments is a major problem. Not all users subscribe to the relevant email distribution lists at mailbase (<http://www.mailbase.ac.uk>), there are problems with cross-posting and managing these lists can also be problematic as email addresses constantly change. Similarly, not all users visit the 'Latest News' section of the MIDAS Web site on a regular basis. One possible solution is the use of push channel technology as an automated method of keeping users notified about new developments. A variety of different products are now available for delivering customised news feeds to users. However, use of push channel technology is still relatively underdeveloped in the UK academic community despite its potential.

Another emerging technology is the use of plugin and helper applications which embed additional functionality within Web browsers and also enable users to access files in more specialised formats. For example, the Adobe Acrobat Reader plugin enables users to view, navigate and print documents, such as electronic journal articles, held in PDF format. From the MIDAS perspective, one of the key benefits of plugin technology is the extent to which it can significantly reduce interface development times. However, less technically competent users do express concerns about the concept of having to locate, download and install plugins – although in the future these tasks may be performed automatically by the browser.

MIDAS has also been experimenting with use of other types of plugin. As part of an ESRC funded project, MIDAS has been developing an entirely Web based interface to 1991 Census area statistics. A major component of this interface is the map based front end to the 1991 Census area statistics database. By downloading and installing the Autodesk MapGuide Viewer (<http://www.mapguide.com/>) the user is able to embed limited desktop mapping functionality within the Web browser. This enables the user to interact dynamically with a multi-layered map which incorporates digital map data from the 1:1,000,000 Bartholomew Europe dataset and digitised 1991 Census output area boundaries. By panning and zooming the user is able to identify and select areas for which 1991 Census area statistics are to be extracted.

Accessibility for all?

It has become increasingly apparent that whilst new digital telecommunications technologies can significantly widen access to data and information resources on a world wide basis they can also serve to reduce accessibility for other groups. It is in that context that many organisations such as the World Wide Web Consortium (W3C) are looking at ways in which access and usability can be improved for people with disabilities (<http://www.w3.org/WAI/>) or with access to text only browsers. It is clear from the ongoing debate on accessibility and usability issues as part of the design of HTML 4.0 (<http://www.w3.org/WAI/References/HTML4-access/>) that it will be increasingly important for data and information service providers, such as MIDAS, to look at the way in which documents are structured and whether there is an over reliance on graphics for information presentation (e.g. frames), navigation aids and/or resource discovery (e.g. clickable image maps). Graphics and multimedia have an important role to play but this should not be at the expense of textual content and/or network efficiency. However, in an era of rapid technological development there will always be a tension between the desire to exploit leading edge technology in interface development and the requirements of low technology users and/or those with special needs.

Conclusion

The Web has already made a tremendous impact on the way in which data centres, such as MIDAS, deliver data and information to the desktop. In order to respond to changing user expectations it will be strategically important for MIDAS to continue to develop more Web based interfaces to its services – although these might not represent the only access method. However, for certain datasets this may not be achievable in the short term due to the absence of Web enabled versions of all the key data access/analysis packages used by the service. Therefore, for certain data resources it may be necessary for MIDAS to develop its own Web based interfaces – which may require a restructuring of the underlying data formats.

Irrespective of the access method, the Web does offer considerable scope for promoting more widespread and effective use of the unrivalled data and information resources that have been made available UK academic community. In this context, it is vitally important that the JISC funded data centres and data services continue to develop and enhance the knowledge base relating to particular data and information resources and provide access to more didactic materials. This will facilitate more effective resource discovery as well as contributing to the development of virtual learning environments and hopefully start to narrow the gap between actual and potential use of the DNER in teaching and research.

Acknowledgements

Keith Cole and Anne McCombe both of whom work for MIDAS have made significant contributions to this paper. Keith Cole as well as being a manager of the MIDAS services is also Director of the ESRC/JISC 1991 Census Dissemination Unit and the Assistant Director (Technical) of the Cathie Marsh Centre for Census and Survey Research (CCSR) at the University of Manchester. Keith has presented a similar paper specifically geared to Social Scientists at IRISS '98 in the UK.

References

MIDAS (<http://www.mimas.ac.uk/>)

Joint Information Systems Committee (<http://www.jisc.ac.uk>)

Distributed National Electronic Resource (DNER) (http://www.jisc.ac.uk/cei/dner_colpol.html).

COPAC (<http://irwell.mimas.ac.uk/sj/>)

SuperJournal Project (<http://www.superjournal.ac.uk/sj/>).

NetEc (<http://netec.mcc.ac.uk>)

GENUKI (<http://www.genuki.org.uk>)

Office for National Statistics Databank (<http://www.mimas.ac.uk/ons/>)

SAS/IntrNet software product (<http://www.sas.com/software/components/intrnet.html>)

ESRI Map Objects Internet Map Server (<http://www.esri.com>)

KINDS (<http://www.kinds.ac.uk/kinds/>)

MIMAS: The Spatial Side (<http://www.kinds.ac.uk/kinds>)

Surpop V 2.0: Introduction (<http://census.ac.uk/cdu/surpop>)

Mailbase (<http://www.mailbase.ac.uk>)

Autodesk MapGuide (<http://www.mapguide.com>)

Web Accessibility Initiative (WAI) (<http://www.w3.org/WAI>)

WAI Resource : HTML 4.0 Accessibility Improvements
(<http://www.w3.org/WAI/References/HTML4-access.html>)

* Paper presented at the 1998 IASSIST/CSS Conference,
Yale University, New Haven, USA, May 19-22, 1998, Julia
Chruszcz, Keith Cole and Anne McCombe, University of
Manchester, Manchester Computing, Oxford Road,
Manchester M13 9PL.

The History Data Service – Using Technology to Enhance Access

The History Data Service (HDS) <<http://hds.essex.ac.uk>> is funded by the UK Joint Information Systems Committee (JISC) <<http://www.jisc.ac.uk>> to collect, manage, and encourage re-use of digital resources which result from or support historical research and teaching. The HDS is located and integrated in the UK Data Archive <[\[dawwww.essex.ac.uk\]\(http://dawwww.essex.ac.uk\)> and is the Arts and Humanities Data Service \(AHDS\) <<http://ahds.ac.uk>> service provider for the historical disciplines. The AHDS provides archival, training and other functions to the archaeology, history, performing arts, textual studies and visual arts communities, and consists of five subject-based service providers and a managing executive. The HDS collection covers a time period from the late tenth century to the mid twentieth century, and includes a wide range of historical data, which has been transcribed or compiled from original sources.](http://</p></div><div data-bbox=)

The HDS is committed to using technology to improve access to its collection through a programme of work that is essentially *needs* rather than *technology*-driven. Although this programme of work aims to make effective use of established and state of the art technologies and concepts for data and metadata storage, presentation and delivery, the real emphasis is on recognising and responding to end-users' needs. The HDS has an active and ongoing policy of consulting with actual and potential users. For example, in April 1998 the HDS held a workshop <http://hds.essex.ac.uk/reports/user_needs/final_report01.stm> to explore, assess and prioritise the needs of end-users in the historical community.

The central goal of this programme of work is to improve and increase access to historical data by making the location, identification, assessment and use of data easier. The objective is to get more people using and experimenting with historical data, and the HDS is aiming to gradually extend its user-base into the less computer-literate and non-computing segments of the historical community by acting as an information provider as well as a data provider. In a very broad sense, the HDS is seeking to improve the relationship between end-users and data and it is essential that this is carried out within existing resources.

by Cressida Chappell, Oscar Struijvé, Sheila Anderson *

The HDS approach to using technology to enhance access is underpinned by a model called TESTMIX, short for telescope, stethoscope and microscope. The principle behind this model is that users want to locate, identify and assess data. The model maps user needs and actions to system functions, and it maps the system functions to the system

components or tools required by users.

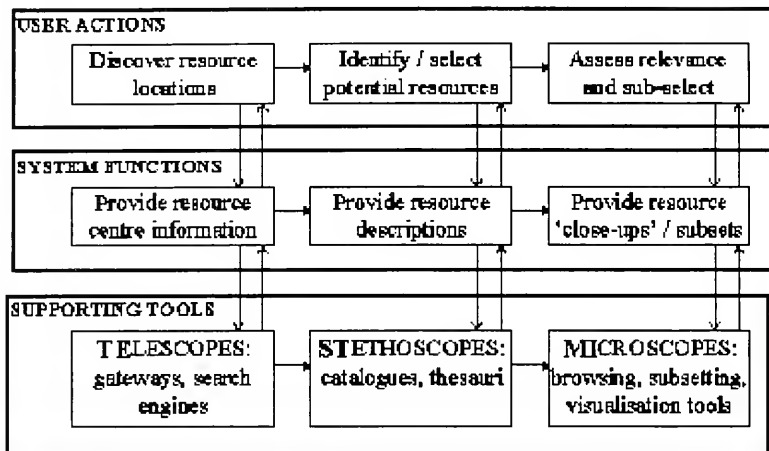
The top level of this model deals with user actions and needs. From left to right we have the need to discover information about data centres, the need to identify and select suitable data, and lastly the need to make detailed assessments of the relevance of data, along with the need to explore data in detail and make sub-selections. The next layer maps user needs and actions to system functions, so we have the provision of information about data centres, the provision of searchable data descriptions, and lastly the provision of data 'close-ups' and subsets. The next layer maps the system functions to the system components or supporting tools that are required, so we have gateways and search engines likened to telescopes; catalogues and thesauri likened to stethoscopes; and browsing, subsetting and visualisation tools likened to microscopes.

The HDS uses TESTMIX as a framework to structure and prioritise development work; to clarify relationships between services and systems; to identify scope for improvement and collaboration; and as a focus point for technical and service activities. It stimulates a coherent approach, and provides simple metaphors for communication. Within this framework the HDS is implementing a multilevelled strategy to improve and increase access to data.

Increasing the number of metadata access points

The first level involves increasing the number of metadata access points and the HDS has two different approaches. One approach uses conventional online catalogues and operational examples include the UK Data Archive's information retrieval system, BIRON <<http://biron.essex.ac.uk/cgi-bin/biron>> and the CESSDA IDC (Council of European Social Science Data Archives Integrated Data Catalogue) <<http://dastar.essex.ac.uk/Cessda/IDC>> Information about the HDS collection is

TESTMIx model



also being made available through the prototype AHDS Integrated Access Gateway <http://prospero.ahds.ac.uk:8080/ahds_live> which is based upon the Dublin Core and the Z39.50 network applications protocol, and which acts as a virtual union catalogue for the collections of the five subject-based service providers. In the coming months information about the HDS collection will also be accessible via the Cheshire Information Retrieval System – an SGML-based system which utilises the Data Documentation Initiative (DDI) Codebook DTD.

The other approach uses a tree-based structure as an alternative way of accessing information about the HDS collection. This will allow users to adopt a 'drill down' approach to locating data in addition to the more sophisticated search options offered by online catalogues. Users will be able to drill down to metadata via three dimensions. A time dimension by centuries, a geographic dimension by countries and administrative subdivisions within the UK, and a subject categories dimension.

Providing online access to additional information

The second level involves providing online access to additional information and will allow users to access to types of information that are not generally found in

catalogue records. In particular we are interested in providing users with access to online documentation with the option to preview a sample of data. We believe that this will make it much easier for users to make detailed assessments of the suitability of data and that it is a more efficient way of supplying users with information.

These services will initially apply to areas of the HDS collection where there is critical mass of related materials, because there is a greater potential to create additional documentation. We are intending to apply this approach to a collection of early twentieth century surveys, a collection of European state finance data, and a collection of electoral poll book data. These services will be freely available to all users and will not require registration.

Providing a data and documentation ftp service

The third level involves providing a data and documentation ftp service and will give registered users online access to the vast majority of the HDS collection. We envisage that users who register with the HDS will be able to select and download data when they require it in suitable easy-to-use formats such as tab or comma delimited ASCII. The main exceptions will be difficult to use data and the minority of HDS data which has more

restrictive access conditions.

Developing online browsing, sub-setting, combining and downloading facilities

The fourth level involves the development of online browsing, sub-setting, combining and downloading facilities for major collections of value-added data and will allow registered users to explore fully documented data collections online. The HDS has developed this service for a large collection of nineteenth and twentieth century statistics, the Great Britain Historical Database (GBHD), and work is now being carried out on developing a similar service for a large collection of individual-level non-anonymised British historical census data which includes the 1881 Census for England and Wales digitised by the Genealogical Society of Utah and the UK Federation of Family History Societies.

The GBHD has been assembled by Humphrey Southall at Queen Mary Westfield College, London and incorporates demographic statistics, marriage statistics, mortality statistics, employment statistics, trade union statistics, government unemployment statistics, poor law statistics and small debt statistics. The GBHD Online system allows

users to sub-set this data by geographical area, at present either standard regions and/or counties. The system also allows users to specify the tables to be searched for relevant data and to specify the variables to be included in the result. Finally the resulting subset and customised documentation can be viewed and browsed online or downloaded to the user's workstation by FTP. Data is formatted as fixed width ASCII with headers, which can be imported into a variety of software packages for further manipulation and analysis. For more information about GBHD Online (including details about registering as a user) please see the GBHD Online webpages at <<http://hds.essex.ac.uk/gbh.stm>>

The HDS is confident that this multilevelled strategy will encourage and enhance use of and experimentation with the HDS collection. We believe that this programme of work will widen and improve access to historical data, and that it will contribute to the creation of an infrastructure which will enable historians and others to explore the full potential offered by digital resources.

*Paper presented at the 1998 IASSIST/CSS Conference, Yale University, New Haven, USA, May 19-22, 1998, by C.Chappell, O. Struijvé, S. Anderson, History Data Service, UK Data Archive, University of Essex



INTERNATIONAL ASSOCIATION FOR
SOCIAL SCIENCE INFORMATION
SERVICE AND TECHNOLOGY

• • • • •
ASSOCIATION INTERNATIONALE POUR
LES SERVICES ET TECHNIQUES
D'INFORMATION EN SCIENCES
SOCIALES

The **International Association for Social Science Information Services and Technology (IASSIST)** is an international association of individuals who are engaged in the acquisition, processing, maintenance, and distribution of machine readable text and/or numeric social science data. The membership includes information system specialists, data base librarians or administrators, archivists, researchers, programmers, and managers. Their range of interests encompasses hard copy as well as machine readable data

Paid-up members enjoy voting rights and receive the **IASSIST QUARTERLY**. They also benefit from reduced fees for attendance at regional

and international conferences sponsored by IASSIST.

Membership fees are:

Regular Membership: \$40.00
per calendar year.
Student Membership: \$20.00
per calendar year.

Institutional subscriptions to the quarterly are available, but do not confer voting rights or other membership benefits.

Institutional Subscription:
\$70.00 per calendar year
(includes one volume of the
Quarterly)

Membership form

I would like to become a member of IASSIST. Please see my choice below:

Options for payment in Canadian Dollars and by Major Credit Card are available. See the following web site for details:

<http://datalib.library.ualberta.ca/iassist/mbrship2.html>

- ☐ \$40 (US) Regular Member
- ☐ \$20 Student Member
- ☐ \$70 Subscription (payment must be made in US\$)
- ☐ List me in the membership directory
- ☐ Add me to the IASSIST listserv

Please make checks payable, in US funds, to IASSIST and Mail to:

**IASSIST,
Assistant Treasurer
JoAnn Dionne
50360 Warren Road
Canton, MI 48187
USA**

Name: _____

Job Title: _____

Organization: _____

Address: _____

City: _____

State/Province: _____

Postal Code: _____

Country: _____

Phone: _____

FAX: _____

E-mail: _____

URL: _____

Return Undelivered Mail To:

IASSIST QUARTERLY

c/o Wendy Treadwell
1758 Pascal St. North
Falcon Heights, MN 55113
USA

OCT 28
1994

Serials Department/SERLIBS82186344)
Univ of North Carolina-Chapel Hill
CB #3938 Davis Library
Chapel Hill NC 27514-8890
U.S.A.